

Avaliação Formativa em Psicologia: Instrumento para Análise de Material Instrucional

Formative Evaluation in Psychology:
Instrument for Instructional Material Analysis

Evaluación Formativa en Psicología:
Instrumento para Análisis de Material Instrucional

**Fabiana Queiroga,
Jairo Eduardo Borges-
Andrade & Rodrigo
Araújo de Miranda**

Universidade
de Brasília

Artigo



Resumo: Livros de texto para o ensino superior muito raramente são objeto de avaliação formativa. Este estudo teve como objetivo desenvolver e validar para o contexto brasileiro um instrumento para a captação de dados de avaliação formativa, visto que no Brasil ainda são escassas as tentativas de aplicação de avaliações dessa natureza. Contou-se com a participação de duas turmas de estudantes da graduação do curso de Psicologia e duas do curso de Administração de uma universidade pública federal. Após ser submetido a uma análise fatorial exploratória, o instrumento apresentou indicadores psicométricos que apontaram evidências de validade de construto, sendo útil para a avaliação de materiais instrucionais. Espera-se que este estudo possa contribuir para o fortalecimento do processo de avaliação formativa de materiais didáticos no contexto brasileiro.

Palavras-chave: Qualidade da educação. Ensino em Psicologia. Materiais didáticos.

Abstract: Text books for undergraduation are seldom the focus of formative evaluation. The goal of this study was to develop and validate for the Brazilian context an instrument for data collection in formative evaluation. Brazil has scarce attempts of this nature. Participants of the study were psychology and management undergraduate students in a federal public university. Data collected with these participants were submitted to an exploratory factor analysis and the results have pointed out evidences of construct validity. The instrument may be useful for the evaluation of instructional materials. This study may contribute for the strengthening of the process of formative evaluation of teaching materials.

Keywords: Educational quality. Psychology teaching. Teaching materials.

Resumen: Libros de texto para la enseñanza superior muy raramente son objeto de evaluación formativa. Este estudio tuvo como objetivo desarrollar y validar para el contexto brasileño un instrumento para la captación de datos de evaluación formativa, visto que en Brasil aún son escasos los intentos de aplicación de evaluaciones de esa naturaleza. Se contó con la participación de dos grupos de estudiantes de la graduación del curso de Psicología y dos del curso de Administración de una universidad pública federal. Después de ser sometido a un análisis factorial exploratorio, el instrumento presentó indicadores psicométricos que señalaron evidencias de validez de constructo, siendo útil para la evaluación de materiales de instrucción. Se espera que este estudio pueda contribuir con el fortalecimiento del proceso de evaluación formativa de materiales didáticos en el contexto brasileño.

Palabras clave: Calidad de la educación. Enseñanza en Psicología. Materiales didáticos.

A qualidade de livros de texto para o ensino superior geralmente é avaliada por meio de especialistas em conteúdo, quando é. Esses livros são concebidos como parte fundamental de estratégias instrucionais que visam a promover processos de aprendizagem entre estudantes. No entanto, naquela avaliação de qualidade, não são colhidos dados com esses estudantes. O presente texto descreve uma experiência de avaliação dessa natureza e o instrumento que foi desenvolvido para viabilizá-la.

As estratégias instrucionais envolvem a exposição sistemática de aprendizagens a experiências que os ajudarão a adquirir informações verbais, estabelecerem estratégias cognitivas ou desenvolverem habilidades intelectuais, motoras ou novas atitudes (Gagné, 1987). Segundo o autor,

a realização do planejamento instrucional eficiente deve ter como meta a aprendizagem completa. O mencionado planejamento deve ser pautado no levantamento das necessidades dos aprendizes (o que precisa ser aprendido?) e no tipo de resultado de aprendizagem esperado. As estratégias devem ser elaboradas antes de os materiais instrucionais serem preparados ou selecionados, e devem incluir: a) métodos, materiais pretendidos e definição do público-alvo; b) os meios instrucionais que serão utilizados; c) o local de realização do evento e d) os métodos de integração desses elementos.

De acordo com a classificação de Jonassen, Grabinger e Harris (1990), há dois tipos de estratégia instrucional. A *estratégia macroinstrucional* é um plano mais geral de experiência de aprendizado, com foco mais

amplo (cursos ou módulos). É a maneira que os planejadores instrucionais têm de ajudar os aprendizes a atingir o objetivo geral do curso. A *estratégia microinstrucional* é um plano específico que engloba cada parte da experiência de aprendizado, como um capítulo/unidade de um curso ou módulo. É a maneira que os planejadores instrucionais têm de ajudar os aprendizes a atingir os objetivos específicos do curso. As estratégias instrucionais não devem ser confundidas com estratégias de aprendizagem, uma vez que estas representam qualquer ação que os aprendizes utilizam para codificar, armazenar, acessar, organizar, compreender e transferir informações com mais eficiência (Pantoja & Borges-Andrade, 2009).

A retroalimentação do sistema instrucional configura uma etapa importante para o sucesso da aprendizagem na medida em que esse procedimento permite a obtenção de informações sobre a eficácia das estratégias instrucionais planejadas (Gagné, 1988). É o momento em que o planejador pode reformular aquilo que não proporcionou o alcance esperado dos objetivos e confirmar as estratégias que funcionaram bem. A avaliação de um curso é definida como qualquer tentativa cujo objetivo é obter informações sobre os efeitos desse curso e determinar o seu valor a partir dessas informações (Borges-Andrade, 2006). Tal definição pressupõe investigações antes, durante e após o curso. Para se determinar esses efeitos, é necessário que se conheça a situação antes do evento e que ela seja comparada com a situação depois do curso.

Evidências em um empreendimento cujo propósito é a avaliação do programa instrucional devem procurar responder as seguintes questões específicas a lições, tópicos, cursos ou sistemas instrucionais: a) em que medida os objetivos específicos propostos na instrução foram atingidos?, b) de quais maneiras e em qual estágio a

unidade nova é melhor que a antiga?, e c) houve possíveis efeitos adicionais, quais foram e em que medida são melhores ou piores que as unidades anteriores? A questão que o processo de avaliação deve responder pode ser dividida em duas áreas diferentes de tomadas de decisão. Tem sido comum, desde o artigo de Scriven, Tyler e Gagné (1967) que virou uma referência fundamental sobre o assunto, definir essas duas áreas como avaliação formativa e avaliação somativa.

A avaliação somativa ocorre, normalmente, após a finalização do programa instrucional. Seu objetivo é permitir a elaboração de conclusões sobre o funcionamento do modelo instrucional a fim de prover dados para a tomada de decisão sobre a adoção do modelo de aprendizagem utilizado. Em geral, a avaliação somativa preocupa-se com a efetividade de um sistema instrucional, curso ou tópico (Borges-Andrade, 2006). As lições individuais são tomadas como componentes das unidades, porém raramente são analisadas em separado. O processo de avaliação somativa está preocupado, antes de tudo, com os resultados obtidos por meio do aprendizado, que são conhecidos a partir de observações realizadas quanto aos testes aplicados de acordo com os objetivos especificados na instrução. É denominada avaliação somativa porque através dela se pretende obter evidências sobre os efeitos somados em um conjunto de lições, originando assim uma unidade ou sistema instrucional mais abrangente. Essas evidências podem incluir informações sobre os pontos positivos ou negativos alcançados em uma lição particular, e tal fato pode ser usado em uma avaliação formativa ou em uma revisão posterior. No Brasil, a avaliação somativa mais clássica é a do SAEB (Sistema de Avaliação da Educação Básica), que é realizada desde 1995 com os estudantes do final dos ciclos do ensino fundamental (4ª série ou 5º ano) e do ensino médio (8ª série ou 9º ano).

Segundo Bell e Cowie (2001b), há a tendência de os estudantes acreditarem que os materiais estão corretos e são efetivos. Quando a aprendizagem não ocorre satisfatoriamente, tendem a atribuir a culpa pelo insucesso a si mesmos ou ao professor.

A avaliação formativa, por outro lado, pode ser definida como um processo utilizado por professores e aprendizes para reconhecer e responder ao aprendizado ao longo do processo de consolidação desse aprendizado, conforme a opinião de Bell e Cowie (2001a); portanto, segundo esses autores, a avaliação formativa tem por objetivo coletar dados sobre um programa instrucional junto à clientela-alvo enquanto o programa está sendo executado, com a finalidade de se elaborar modificações para adequá-lo às necessidades da clientela. Ao final de cada capítulo ou módulo do programa instrucional, os estudantes respondem um questionário estruturado que tem por objetivo mensurar a adequação e o seu nível de satisfação em relação ao material utilizado. Esse modo de avaliação visa à tomada de decisões sobre como revisar um programa instrucional ao longo do curso, ou seja, as evidências são coletadas, interpretadas e utilizadas para melhorar o próprio programa instrucional.

As avaliações formativas são, normalmente, caracterizadas pela informalidade. O modelo instrucional – lição, tópico, curso ou sistema instrucional – é testado e muitas observações são realizadas, ao mesmo tempo, por professores ou por outro observador, e devem ter como foco o desempenho dos estudantes durante as tarefas. O que for observado pelos professores deve ser incluído no processo de avaliação formativa após a realização das tarefas. A informalidade nos procedimentos de coleta de dados não pode influenciar na precisão dos dados, ou seja, dados quantitativos são necessários em avaliação formativa (Bell & Cowie, 2001b). Ainda segundo esses autores, a abordagem metodológica utilizada em avaliação formativa difere entre os casos, ou seja, cada objetivo em estudo exige métodos particulares de coleta de dados, dependendo das condições locais; portanto, não há padronização dos métodos. Objetivando a tomada de decisões sobre a eficiência e

a viabilidade de um programa instrucional, um ponto inicial que se repete em ocasiões posteriores é a revisão do conteúdo do programa instrucional por um sujeito com *expertise* no campo abordado. A partir dos dados coletados por meio de observações, questionários e testes, são tomadas decisões sobre a possibilidade de o conteúdo estudado ser mantido no programa instrucional, reformulado ou até mesmo descartado.

As avaliações formativa e somativa requerem métodos científicos cuidadosos, a fim de se garantir a fidedignidade das evidências obtidas. No caso da avaliação formativa, deve-se tomar algumas precauções na sua realização. A principal delas trata da explicação dada aos estudantes sobre o processo de avaliação de determinados materiais instrucionais, uma vez que eles não estão acostumados a criticarem, construtivamente, tais materiais. Segundo Bell e Cowie (2001b), há a tendência de os estudantes acreditarem que os materiais estão corretos e são efetivos. Quando a aprendizagem não ocorre satisfatoriamente, tendem a atribuir a culpa pelo insucesso a si mesmos ou ao professor. É necessário estabelecer um bom *rapport* e incentivá-los a fazer críticas construtivas aos materiais instrucionais avaliados. Eles precisam entender a importância de sua participação no processo de avaliação formativa e os ganhos que poderão ser obtidos com os resultados.

Em alguns países, é comum a prática da avaliação formativa como ferramenta para assimilação do conteúdo em uma disciplina escolar, permitindo a constante troca de idéias e de *feedback* entre estudantes e professores, o que facilita o processo de aprendizagem. Um livro clássico da área, escrito por Dick e Carey (1978), propõe três estágios principais em uma avaliação formativa, a saber: avaliação um a um, avaliação de uma amostra e avaliação de campo. No primeiro estágio, o professor seleciona dois ou mais

estudantes que tenham médias parecidas (é preferível que metade dos escolhidos tenha notas um pouco acima da média e a outra metade apresente notas abaixo da média) e explica que o objetivo do estudo é avaliar a reação dos estudantes a um novo material instrucional.

O segundo estágio, o de avaliação de uma amostra, é caracterizado pela escolha aleatória de um grupo de 10 a 20 estudantes com particularidades diferentes a fim de comparar tais dados com os obtidos ao longo do primeiro estágio. A orientação quanto ao objetivo do estudo nesse segundo estágio difere em relação ao anterior; o professor explica que o material apresentado se encontra na fase de elaboração e que se está buscando o *feedback* dos estudantes sobre o funcionamento desse material. Alguns métodos também podem ser utilizados na avaliação de uma amostra, tais como: aplicação de um questionário, com perguntas do tipo: *you understand what is expected of you at the end of the course?, were included enough practical exercises?*, entre outras. Podem ser realizados grupos focais com alguns estudantes da amostra a fim de se obter outros dados.

No último estágio, escolhe-se um grupo de aproximadamente 30 estudantes que representem uma amostra significativa do público-alvo a quem será destinado o novo material instrucional. Dessa forma, todas as ferramentas necessárias à aplicação do material devem ser revisadas previamente a essa etapa. O professor deve procurar simular uma situação de aprendizagem mais próxima possível daquela para a qual será destinado o material instrucional. Algumas precauções metodológicas são necessárias na aplicação de tal estágio, principalmente em relação ao “efeito halo”, ou seja, obter a participação

dos estudantes somente porque o contexto instrucional foi alterado.

No Brasil, não foram encontrados relatos da utilização de avaliação formativa no processo de elaboração de livros que têm finalidades instrucionais em Psicologia. Ela é mais utilizada para o desenvolvimento de materiais de cursos à distância veiculados pela *web* (ver, por exemplo, os relatos de experiências sistematizados por Otsuka & Rocha, 2002). O presente estudo tem como objetivos analisar os resultados obtidos com a avaliação formativa dos capítulos de um livro na subárea da Psicologia Organizacional e do Trabalho (POT) e validar instrumento de avaliação formativa. Espera-se contribuir, com uma base metodológica mínima, para o fortalecimento do processo de avaliação formativa, visando a torná-lo uma prática de revisão de materiais instrucionais em Psicologia e em outras áreas do conhecimento. Lamentavelmente, mesmo as melhores editoras, ao receberem propostas de textos para publicação no formato de livros didáticos de cursos superiores, limitam o processo de avaliação a uma apreciação realizada por pares ou especialistas em conteúdo. A opinião e a aprendizagem dos estudantes desses cursos, que serão os usuários daquele material, parecem não ser importantes como dados empíricos a serem levados em conta nesse processo de avaliação.

Método

Amostra

O material utilizado no curso consistiu em versões preliminares de capítulos do livro posteriormente publicado por Zanelli, Borges-Andrade e Bastos (2004). Essas versões, escritas por vários autores que são docentes em universidades de todo o Brasil, foram utilizadas ao longo de um semestre letivo com estudantes de graduação (maioria entre 3º e

4º semestres) da Universidade de Brasília (UnB) que estavam agrupados em 4 turmas – 2 do Curso de Psicologia e 2 do Curso de Administração. O mencionado livro estava sendo elaborado para atender a esse tipo de clientela e teve objetivos instrucionais previamente formulados em termos de competências, pelos autores dos capítulos, e discutidos intensamente entre eles e com os organizadores do livro. Tais objetivos estão descritos nas partes I, II, III e IV do mencionado livro, quando elas apresentam os capítulos ao leitor. A Tabela 1 exibe o número de respostas obtidas no questionário de avaliação de cada um dos capítulos, de acordo com o curso do estudante.

Tabela 1. Freqüência de resposta dos estudantes (N = 507).

Capítulos	Curso	
	Psicologia	Administração
O mundo do trabalho	38	23
Conceito e perspectivas de estudo das organizações	33	-
Dimensões básicas de análises	-	27
A Psicologia e a produção de conhecimento em organizações e trabalho	38	-
Cognição nas organizações de trabalho	20	18
Aprendizagem humana em organizações e trabalho	-	23
Motivação no trabalho	9	13
Emoções e afetos no trabalho	21	12
Saúde mental e trabalho	17	-
Vínculos do indivíduo com a organização e com o trabalho	24	11
Socialização organizacional	17	17
Grupos e equipes de trabalho nas organizações	25	14
Poder nas organizações	12	11
Cultura organizacional	19	-
Diversidade cultural	-	32
Inserção profissional do psicólogo em organizações e no trabalho	33	-
Total	306	201

Conforme pode ser observado, maior quantidade de avaliações foi realizada por estudantes do Curso de Psicologia. Dos 16 capítulos avaliados, 4 foram avaliados apenas pelos estudantes de Psicologia e 3 apenas pelos estudantes do Curso de Administração.

Material avaliado

Os 16 capítulos do livro de Zanelli, Borges-Andrade e Bastos estão agrupados em 4 partes:

Parte I – O contexto: contempla três capítulos que abordam temas como o mundo do trabalho, os conceitos e perspectivas e as dimensões básicas de análise das organizações;

Parte II – O indivíduo no contexto: aborda os assuntos sobre motivação no trabalho, cognição, emoções e afetos, aprendizagem humana em O&T, saúde mental e vínculos do indivíduo com O&T, totalizando, portanto, seis capítulos;

Parte III – Processos organizacionais: é composta por cinco capítulos, que incluem os temas de socialização e cultura organizacional, grupos e equipes, poder nas organizações e diversidade no contexto organizacional;

Parte IV – Atuação profissional e produção do conhecimento: agrupa os dois últimos capítulos

do material e aborda os temas de inserção profissional do psicólogo em O&T e a relação da Psicologia com a produção do conhecimento em O&T.

Instrumento de avaliação

Para cada capítulo avaliado, os estudantes responderam um questionário composto por 30 itens, utilizando uma escala de grau de concordância do tipo *likert* de 10 pontos que variava de 1 = *discordo totalmente* a 10 = *concordo totalmente*. Os itens foram formulados visando a cobrir os aspectos básicos propostos por Gagné e Briggs (1974) para avaliação formativa de material instrucional e foram agrupados conforme é descrito a seguir:

1. Objetivo: foram elaborados quatro itens para avaliar aspectos relacionados aos objetivos do capítulo, tais como clareza, relevância e pertinência;
2. Conteúdo: para avaliar o conteúdo do capítulo, foram elaborados nove itens que questionavam a opinião do estudante sobre a coerência entre os assuntos apresentados, a relação entre os assuntos e os objetivos do capítulo e ainda sobre a linguagem e o vocabulário utilizados;
3. Exemplos: quatro itens avaliaram a clareza, a pertinência e a quantidade dos exemplos utilizados no capítulo
4. Ilustrações: de maneira semelhante, cinco itens foram elaborados para avaliar a pertinência das ilustrações para a compreensão do texto, a clareza e a quantidade das ilustrações no capítulo.
5. Estudos e pesquisa: além dos aspectos de pertinência, clareza e quantidade, seis itens foram elaborados para avaliar, ainda, se a apresentação dos estudos ou de pesquisa favoreceram a compreensão de como o tema

abordado no capítulo é estudado no Brasil;

6. Prova: também foram elaborados dois itens sobre um dos materiais utilizados para avaliar o grau de aprendizagem do capítulo. O estudante indicava em que medida a prova havia mensurado os objetivos do capítulo avaliado e em que medida ela cobria os conteúdos abordados.

Procedimento de coleta de dados

Ao longo de um semestre letivo na UnB, os estudantes dos Cursos de Psicologia e de Administração, respectivamente matriculados nas disciplinas Psicologia das Organizações e Psicologia Aplicada à Administração, tiveram aula com professores que adotaram procedimentos padronizados entre si. Essas disciplinas são obrigatórias no currículo mínimo daqueles dois cursos. As turmas de Psicologia ficaram sob a responsabilidade de um mesmo professor, ao passo que as turmas de Administração tiveram dois professores diferentes, que eram estudantes de pós-graduação em estágio de docência, sob a supervisão do professor das turmas de Psicologia das Organizações. Entretanto, com a finalidade de coletar dados de maneira sistemática, os três professores acordaram em adotar as mesmas atividades e materiais instrucionais ao ministrar suas aulas.

Para todas as turmas, as disciplinas tiveram 90 horas aula e ocorreram em 3 dias intercalados da semana, conforme é previsto em suas ementas e é semestralmente programado na UnB; portanto, não ocorreram alterações na maneira como aquelas disciplinas são oferecidas nessa universidade. Foi disponibilizada uma versão eletrônica dos capítulos do material que seria avaliado e avisado aos estudantes que aquela era uma versão quase final do material didático. Em seguida, foram esclarecidos os procedimentos de condução da disciplina para aquele semestre, deixando

claro que os estudantes participariam de uma experiência de avaliação de um livro já aceito para publicação, mas que seus autores e a sua editora ainda acatariam sugestões de alterações. Por meio de um mesmo conjunto de transparências, elaboradas especialmente para cobrir os tópicos fundamentais dos capítulos do futuro livro, os estudantes das quatro turmas assistiram aulas expositivas sobre os capítulos do material instrucional, de modo que a aula seguinte era reservada para a discussão do conteúdo do capítulo por meio de estudos dirigidos realizados em grupos. A terceira aula da semana foi destinada à avaliação dos estudantes por meio de prova escrita. Após a realização da prova, estes recebiam o questionário de avaliação do capítulo para responder. Nesse momento, em cada dia de aplicação desse questionário, era informado que o seu preenchimento era voluntário, e que as respostas não teriam qualquer influência na correção das provas escritas da turma ou na atribuição individual de notas, fossem quais fossem as respostas e tivessem os estudantes decidido respondê-las ou não.

Análise dos dados

Os dados foram tabulados no pacote estatístico SPSS (*Statistical Package for Social Science*) e a versão 14.0 foi utilizada para realizar as análises. A verificação da estrutura empírica do instrumento de avaliação foi realizada por meio de uma análise fatorial exploratória, cujos resultados são apresentados na seção a seguir. Posteriormente, a média das respostas dos estudantes nos fatores encontrados foi relacionada com as medidas de desempenho obtidas nas provas idênticas aplicadas simultaneamente em cada turma.

Resultados

Análises dos pressupostos

Existe muita discussão (ver, por exemplo, Costello & Osborne, 2005; Crocker & Algina,

1986; Gorsuch, 1973) sobre o tamanho adequado da amostra para realização da análise fatorial. Levando-se em conta o critério de número de sujeitos (cerca de 16 observações por item) ou das comunalidades (5 itens apresentaram valores entre 0,24 e 0,36, e o restante, entre 0,40 e 0,72) e as cargas fatoriais produzidas (que variaram entre 0,44 e 0,83, com exceção do item 2.1, cuja carga fatorial foi de 0,30), o tamanho da amostra pode ser considerado adequado (Pasquali, 2005; Tabachnick & Fidell, 2001).

Adicionalmente, foram realizadas análises sobre os casos faltosos e as respostas extremadas, visto que a elevada presença de valores extremos e de valores ausentes poderia prejudicar a confiabilidade dos dados. Essas análises indicaram que alguns itens sobre os exemplos, as ilustrações e as pesquisas dos capítulos obtiveram percentuais acima de 5% de casos em branco. Entretanto, esses casos não se devem à falta de resposta dos estudantes, mas à ausência desses elementos nos capítulos, ou seja, foram respostas do tipo “não se aplica”, recodificadas como valores ausentes. Desse modo, optou-se por não fazer qualquer tratamento nesses casos, visto que eles contemplavam características importantes do material avaliado. Com relação à inspeção de valores extremos, foram seguidas as recomendações de Tabachnick e Fidell e identificados 201 casos uni e multivariados. Contudo, conforme foi relatado em apresentação preliminar desses resultados (Queiroga, Borges-Andrade, & Mauro, 2004), a presença desses casos não altera a estrutura final do instrumento; portanto, nas análises empregadas nesse estudo, optou-se por manter todas as respostas coletadas.

Essa decisão de não excluir os casos extremos acabou afetando sensivelmente os índices de assimetria (que variaram de -2,12 a 1,90) e de curtose (que variaram de -1,14 a 6,73). De acordo com Miles e Shevlin (2001), valores acima de |2,0| já requerem atenção do pesquisador, embora ofereçam pouco perigo na distorção dos resultados. Em consequência, gráficos *scatterplots* utilizados para verificar a presença de um importante pressuposto da análise fatorial, que é a linearidade, também ficaram prejudicados.

Esse pressuposto está diretamente relacionado com a matriz de correlações, que é a base da análise fatorial (Laros, 2005; Pasquali, 2005). Apesar de a inspeção gráfica ter apontado alguns problemas nesse requisito, a robustez da matriz de correlações e a observação do contraste esperado na matriz de correlações não encorajaram a adoção de transformações nas variáveis, por isso, optou-se por realizar a análise fatorial com os dados originais.

Análises principais

Além da inspeção da matriz de correlações, foi verificada a possibilidade de agrupamento dos itens por meio dos índices de KMO (0,89), do teste de esfericidade de Bartlett ($\chi^2 = 5354,419$ (435); $p < 0,001$) e do determinante (0,000000412). Os três indicadores apontaram a adequação da fatorabilidade da matriz (Pasquali, 2005; Tabachnick & Fidell, 2001). Para a decisão sobre o número de fatores, foram contrastados os resultados de dois testes – *scree plot* e análise paralela. O *scree plot* apresentado na Figura 1 sugere a presença de cinco fatores.

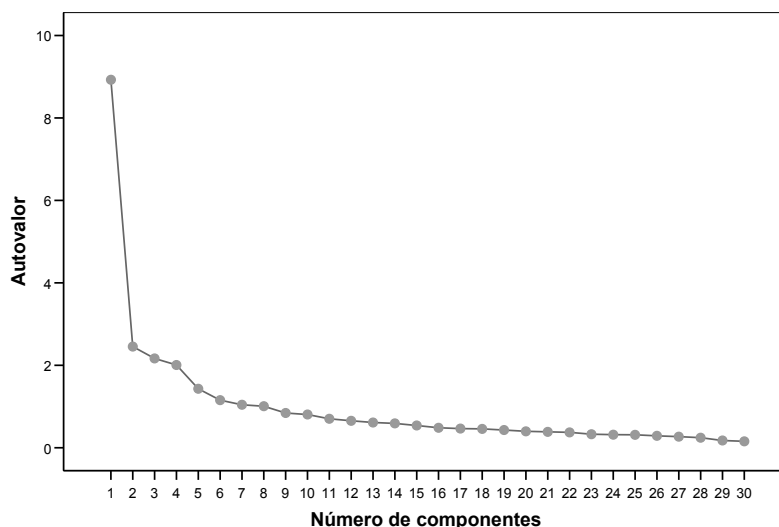


Figura 1. *Scree plot* dos dados obtidos com o instrumento de avaliação do material instrucional.

Entretanto, Zwick e Velicer (1986) recomendam o método da análise paralela de Horn como o mais preciso para essa tarefa: em 92% dos casos, o método indicou o número correto de fatores. A interpretação da análise paralela propõe a retenção de componentes cujos autovalores empíricos são superiores àqueles obtidos paralelamente por meio de matrizes de variáveis randômicas não-correlacionadas, com igual tamanho de amostra e número de variáveis (Laros, 2005). O que essa análise propõe é que os dados empíricos devem explicar mais variância que uma matriz de dados gerada aleatoriamente. Nesse estudo, os autovalores aleatórios foram gerados por meio da sintaxe de Enzmann (2003), e os resultados apontaram a presença de cinco fatores, conforme é mostrado na Tabela 2.

Tabela 2. Análise paralela de Horn – comparação entre autovalores empíricos e aleatórios.

Autovalores	1	2	3	4	5	6	7	8	9	10
Empíricos	8,92	2,45	2,16	2,01	1,43	1,15	1,04	1,01	0,84	0,81
Aleatórios	1,48	1,41	1,36	1,32	1,28	1,25	1,22	1,18	1,16	1,13

Diante da convergência entre os dois critérios utilizados sobre o número de fatores a ser extraído, os 30 itens do instrumento foram submetidos a uma análise fatorial dos eixos principais (PAF), com critério de extração de cinco fatores. O método de redução de dados empregado na PAF tem como base uma matriz de covariância entre as variáveis, ou seja, a variância única e a variância erro são desconsideradas (Laros, 2005). No caso desse estudo, foram realizadas coletas semanais com os mesmos grupos de estudantes, o que gera, portanto, “respostas dependentes” no banco de dados no que tange a esses estudantes, mas não no que concerne ao foco do que avaliavam. Embora não se esteja buscando variáveis latentes no agrupamento de itens desse instrumento, o emprego da PAF ajusta-se melhor a essa situação na medida em que desconsidera a variância erro que foi aumentada com o tratamento dado às respostas dependentes. Por fim, tomou-se a decisão de utilizar a rotação oblíqua do tipo *promax*, em decorrência do grau de interdependência teórica e empírica (as correlações variaram de -0,17 a 0,60) que os fatores estabelecem entre si. A Tabela 3 apresenta a estrutura fatorial proposta para o instrumento.

Tabela 3. Estrutura fatorial do instrumento de avaliação e índices de consistência dos fatores.

Itens	I ($\alpha = 0,86$)	II ($\alpha = 0,85$)	III ($\alpha = 0,83$)	IV ($\alpha = 0,76$)	V ($\alpha = 0,77$)	h ²
2.4 Existe coerência entre os assuntos apresentados	0,80	-0,07	-0,10	0,09	0,05	0,53
1.1 Os objetivos estão claramente formulados	0,73	0,04	0,00	0,01	-0,17	0,51
1.2 Os objetivos foram alcançados com êxito	0,72	-0,15	0,08	0,00	0,09	0,50
1.4 Os objetivos são pertinentes aos conteúdos abordados	0,70	-0,03	-0,03	0,00	0,00	0,46
2.7 A organização do texto consegue manter a motivação para a leitura	0,69	-0,02	0,04	-0,10	0,03	0,55
2.2 As seqüências dos assuntos estão adequadas	0,64	-0,08	0,08	-0,04	0,11	0,47
2.3 Os tópicos dos capítulos estão interligados	0,61	0,00	0,13	0,02	-0,19	0,40
1.3 Os objetivos são relevantes para sua formação profissional	0,59	-0,04	0,03	-0,09	0,09	0,42
2.6 A linguagem utilizada é compreensível	0,59	-0,04	0,03	0,04	0,08	0,36
2.9 O tempo que você teve para leitura foi suficiente para a aprendizagem	0,50	0,10	-0,13	0,04	0,03	0,28
2.1 Diferentes visões sobre o assunto foram apresentadas	0,40	0,17	-0,09	0,18	0,07	0,26
5.5 A quantidade de estudos e pesquisas descritos é suficiente	-0,29	0,78	0,16	0,08	0,05	0,51
5.3 Os estudos favorecem a compreensão de como o tema é estudado no Brasil	0,04	0,73	-0,01	0,02	-0,01	0,56

Continuação da tabela3

5.2 Os estudos estão apresentados em sintonia com o conteúdo do capítulo	0,24	0,65	-0,02	-0,06	0,02	0,68
5.1 Os estudos ou pesquisas estão descritos de maneira clara	0,26	0,63	0,00	-0,03	-0,05	0,66
5.4 Os estudos favorecem a compreensão de como o tema pode ser utilizado	0,17	0,60	-0,12	0,00	0,06	0,48
3.3 A quantidade de exemplos é suficiente	0,11	0,58	0,04	-0,08	-0,02	0,46
3.2 Os exemplos são pertinentes à realidade brasileira	-0,19	0,56	0,05	0,01	0,02	0,24
3.1 Os exemplos estão descritos de maneira clara	0,32	0,45	0,05	-0,02	-0,07	0,50
4.1 As ilustrações estão apresentadas de maneira clara	0,22	-0,10	0,79	-0,03	-0,03	0,72
4.3 As ilustrações estão suficientemente explicadas no texto	0,16	0,00	0,77	0,01	-0,06	0,69
4.2 As ilustrações são pertinentes para a compreensão do texto	0,01	0,09	0,76	0,04	0,00	0,64
4.4 A quantidade de ilustrações é suficiente	-0,27	0,16	0,70	0,03	0,09	0,49
3.4 A quantidade de exemplos é excessiva	0,17	0,00	-0,17	0,81	-0,01	0,65
5.6 A quantidade de estudos e pesquisas descritos é excessiva	0,01	0,08	-0,02	0,76	-0,06	0,57
4.5 A quantidade de ilustrações é excessiva	-0,06	-0,08	0,28	0,68	0,03	0,52
6.2 Esta prova cobriu conteúdo que foi abordado no capítulo	0,04	0,08	-0,01	-0,02	0,75	0,62
6.1 Esta prova mediu objetivos previstos no capítulo	0,17	-0,03	0,02	0,04	0,74	0,65
2.8 A aprendizagem proporcionada pelo capítulo supera os objetivos propostos	0,25	0,16	0,12	-0,07	0,12	0,25
2.5 O vocabulário usado necessita de definições que estão ausentes no capítulo	-0,15	-0,01	0,04	0,23	0,17	0,11

Notas: I = planejamento (11 itens); II = casos (8 itens); III = ilustrações (4 itens); IV = excessos (3 itens); V = avaliação da prova (2 itens). As cargas fatoriais em negrito referem-se aos itens de cada fator. Rotação utilizada: *promax*

Foi considerado um critério de carga fatorial de 0,40 para que o item pertencesse a um fator. Dois itens (2.8 e 2.5) não apresentaram cargas fatoriais em qualquer fator e foram, portanto, desconsiderados. Desse modo, 28 itens agruparam-se em 5 fatores que são descritos a seguir:

Fator I: diz respeito à qualidade dos objetivos dos capítulos do material avaliado bem como ao conteúdo utilizado para o alcance de cada objetivo; portanto, é um fator que diz respeito à avaliação da etapa do *planejamento instrucional*. O item que obteve maior carga fatorial foi 2.4 – Existe coerência entre os assuntos apresentados (0,80), e o que alcançou menor carga foi 2.1 – Diferentes visões sobre o assunto foram apresentadas (0,40). O índice de consistência interna desse fator (alfa de Cronbach) foi considerado bastante satisfatório.

Fator II: refere-se à avaliação das pesquisas e dos exemplos apresentados nos capítulos do livro, e por isso foi rotulado de casos. As cargas fatoriais nesse fator variaram de 0,78 (5.5 – A quantidade

de estudos e pesquisas descritos é suficiente) a 0,45 (3.1 – Os exemplos estão descritos de maneira clara). O alfa de Cronbach desse fator também foi considerado meritório.

Fator III: são agrupados neste fator os itens relacionados à avaliação das ilustrações apresentadas nos capítulos. As cargas fatoriais neste fator variaram de 0,79 (4.1 – As ilustrações estão apresentadas de maneira clara) a 0,70 (4.4 – A quantidade de ilustrações é suficiente), e o índice de consistência interna foi adequado.

Fator IV: neste fator, ficaram agrupados itens que dizem respeito aos aspectos de excessos dos capítulos, ou seja, o quanto os exemplos, as pesquisas e as ilustrações foram apresentadas de maneira excessiva no material. O alfa de Cronbach deste fator indica que sua fidedignidade é moderada, mas não é considerada inaceitável.

Fator V: foram inseridos no questionário itens que procuraram investigar a opinião dos estudantes sobre a adequação da prova referente a cada capítulo, no que concerne à mensuração dos objetivos e à cobertura do conteúdo. Apesar de não contemplar o número mínimo de itens recomendado pela literatura (Costello & Osborne, 2005), este fator abarca os aspectos necessários para a análise da qualidade das provas respondidas pelos estudantes, na opinião destes, e apresenta índice de consistência interna aceitável e excelentes cargas fatoriais (0,74 e 0,75).

Desempenho dos estudantes nas provas e relação com os escores fatoriais

De acordo com o que foi exposto na revisão de literatura, o professor deve associar as medidas de qualidade do material ao desempenho dos estudantes durante as tarefas realizadas no curso. Assim, tomando como base a estrutura fatorial apresentada anteriormente, foi realizado o cálculo das médias nos fatores do questionário, e, posteriormente, essas médias foram comparadas com o desempenho dos estudantes em cada uma das turmas. O desempenho final dos estudantes levou em conta as notas das provas realizadas ao final de cada capítulo assim como a sua performance nos grupos de discussão. Para ambas as atividades (provas e grupo de discussão), havia um gabarito que descrevia o resultado esperado em cada avaliação e que serviu de parâmetro para calcular a pontuação do desempenho dos estudantes. O resultado das notas finais relacionado à avaliação em cada fator do questionário pode ser observado nas Figuras 2 e 3.

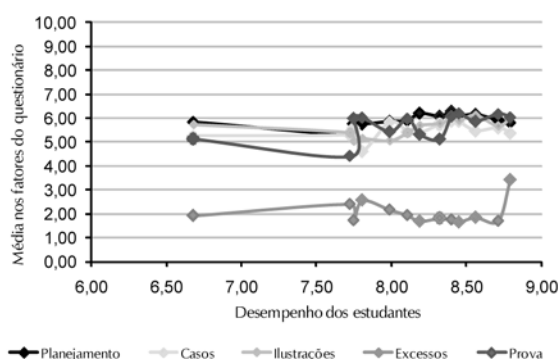


Figura 2. Avaliação dos estudantes do curso de Psicologia.

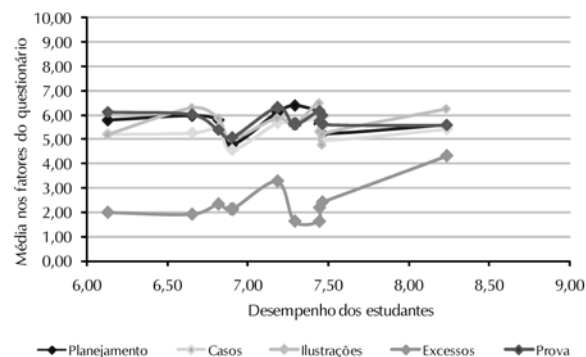


Figura 3. Avaliação dos estudantes do curso de Administração.

Os estudantes de ambos os cursos atribuíram valores medianos para os aspectos dos capítulos avaliados por meio do questionário. Os objetivos e os conteúdos abordados nos capítulos foram os aspectos mais positivos do material na opinião dos estudantes – a média para o fator *planejamento* foi 5,9 para os estudantes da Psicologia, e 5,7 para os estudantes da Administração. O valor modesto observado para o fator *excessos* não indica uma avaliação negativa, mas indica que os estudantes concordam que os exemplos, estudos e ilustrações são suficientes nos capítulos.

Em geral, a avaliação dos aspectos referentes aos capítulos não variou em função do desempenho dos estudantes. No caso do Curso de Psicologia, verificou-se apenas que o grupo de estudantes com notas entre 7,5 e 8,0 avaliou as provas dos capítulos com médias um pouco inferiores. Aqueles com notas acima de 8,5 foram os que indicaram haver mais excessos nos capítulos, visto que a média desse fator foi acima de 3,0.

A avaliação dos estudantes do Curso de Administração também não apresentou grande variação em função do desempenho obtido nos capítulos. Entretanto, aqueles que obtiveram notas entre 6,5 e 7,0 foram os que indicaram as menores médias para os fatores do questionário. Assim como ocorreu no outro curso, os alunos do Curso de Administração indicaram haver poucos excessos dos capítulos do livro, e aqueles cujas notas foram as maiores detectaram mais aspectos em demasia.

Discussão

A avaliação formativa consiste no procedimento de coletar dados sobre um programa instrucional junto à clientela-alvo enquanto o programa está sendo executado, e sua finalidade é a adequação do programa às necessidades da clientela (Bell & Cowie, 2001a). Nesse estudo, procurou-

se avaliar o material instrucional utilizado nos Cursos de Psicologia das Organizações e de Psicologia Aplicada à Administração. Na época, esse material era composto por versões preliminares de capítulos, em fase de revisão final (já previamente validados quanto à propriedade do seu conteúdo, em pelo menos três ocasiões, por pelo menos dois especialistas e por pelo menos um avaliador independente da editora). Para tanto, fez-se necessária a construção e a validação de um instrumento para alcançar de forma adequada esse objetivo.

Para essa construção, foram considerados os aspectos básicos propostos por Gagné e Briggs (1974) para a avaliação formativa de material instrucional. Os resultados alcançados por meio da análise fatorial exploratória indicaram uma estrutura simples de cinco fatores que avaliam os aspectos de planejamento, casos, ilustrações e excessos. Além de apresentarem consistência interna estatisticamente satisfatória (os alfas de Cronbach dos fatores são acima de 0,70), os fatores são nítidos, de fácil interpretação e são condizentes com o que foi previsto teoricamente.

Cabe destacar que os estudantes, ao expressar suas opiniões, foram capazes de distinguir o que era aspecto essencial dos capítulos (Fatores I a IV) do que era a avaliação do seu próprio desempenho (Fator V), o que, de certa maneira, desmistifica um estereótipo prevalente, entre muitos docentes, de que eles não conseguem separar tais opiniões. Da mesma maneira, foram capazes de distinguir o que era estratégia de planejamento de conteúdos (Fator I) e o que era estratégia instrucional para exemplificar e ilustrar tais conteúdos (Fatores II a IV). Ao mesmo tempo, os dados sugerem o quanto são inseparáveis os aspectos de formular, de concatenar e da relevância de conteúdo, naquela estratégia de planejamento (Fator I).

Os exemplos e as ilustrações em excesso (Fator IV) foram mais apontados pelos estudantes que obtiveram escores elevados nas provas, provavelmente porque tais estudantes têm habilidades intelectuais mais desenvolvidas ou estão mais motivados para aprender. A eliminação de exemplos e ilustrações, se fosse realizada de modo sistemático, certamente prejudicaria o aprendizado daqueles que não teriam tais habilidades ou motivação. Como a presente pesquisa trata da validação formativa de um material instrucional impresso, que não tem a flexibilidade de outros materiais de entrega de instrução mais próprios para o ensino individualizado, essa eliminação precisaria ser realizada com extrema cautela.

As poucas diferenças encontradas entre estudantes de Psicologia e de Administração sugerem que o livro é material didático apropriado para clientelas que podem ter diferentes habilidades de entrada. Os estudantes do primeiro curso já tinham cursado outras disciplinas em Psicologia (Social, Geral e Experimental, Psicofisiologia e História e Sistemas, por exemplo). Por outro lado, os alunos de Administração, quando cursaram a disciplina, tinham seu primeiro contato com a Psicologia ou tinham anteriormente cursado só a disciplina Introdução à Psicologia.

De acordo com a estrutura proposta, os dados sobre os capítulos foram analisados e as considerações enviadas para os respectivos autores. Além das avaliações apresentadas nesse estudo, foram enviados aos autores os comentários dos estudantes feitos nas questões abertas do questionário, presentes ao final de cada seção. Assim, apesar de os resultados terem sido positivos, no geral, as análises por capítulo auxiliaram alguns autores a realizar pequenos ajustes, principalmente quanto ao vocabulário utilizado. Por meio do resultado das análises

sobre os excessos, os autores fizeram ajustes na quantidade de exemplos, ilustrações e estudos quando foi pertinente, ou seja, quando havia indícios de desequilíbrios desses elementos nas diferentes seções dos capítulos.

A discussão sobre os resultados deve levar em conta as limitações desse estudo. Um aspecto que pode merecer maior atenção diz respeito à composição da amostra e do procedimento de coleta dos dados. Conforme foi relatado na seção do método, a coleta foi realizada semanalmente com o mesmo grupo de estudantes, de modo que o banco de dados é composto por várias respostas de um mesmo estudante. Essas respostas, no entanto, a cada semana estavam sendo produzidas para avaliar distintos objetos, pois evidentemente os capítulos do livro se sucediam.

Apesar de a análise fatorial exploratória não ser recomendada para o caso de “medidas dependentes”, a estrutura fatorial resultante mostrou-se coerente e bastante confiável. Esse resultado pode ter sido alcançado em virtude do uso da PAF, em que as variâncias erro e específicas são desconsideradas na matriz de correlação utilizada como ponto de partida para a redução dos dados. Nesse sentido, os resultados dos gráficos em que são cruzadas as médias nos fatores e o desempenho dos estudantes nas provas serviram como ferramenta para analisar e interpretar a estrutura do instrumento. Embora os estudantes de Psicologia e Administração tenham apresentado padrões semelhantes nas suas avaliações, as pequenas variações dentro dos grupos contribuíram para representar a coerência da estrutura do instrumento. Não obstante, cabe fazer a ressalva que replicações futuras desse estudo devem levar em consideração esse procedimento de coleta de dados e verificar a estabilidade da estrutura fatorial aqui apresentada.

A análise das médias atribuídas aos fatores do questionário, em função do desempenho dos estudantes nas avaliações, também confirma a coerência da estrutura do instrumento. Desse modo, é possível afirmar que o instrumento proposto apresenta evidência de validade de construto e pode ser útil para a avaliação formativa de materiais instrucionais. Entretanto, é prudente ressaltar que o questionário ora apresentado apenas serviu de instrumento para a avaliação de um material específico em Psicologia Organizacional e do Trabalho, que foi o foco do livro didático adotado em duas disciplinas de graduação. Para a avaliação de outros tipos de materiais instrucionais, poderá ser necessário fazer adaptações para contemplar outras especificidades. O fato de o instrumento ter avaliado um material para o ensino em

Psicologia não limita seu uso a essa área; contudo, é recomendável testar novamente a estrutura do instrumento na avaliação formativa de outros materiais instrucionais, seguindo os passos registrados no presente estudo.

Espera-se que os resultados aqui relatados possam contribuir para a construção de uma base metodológica mais forte do processo de avaliação formativa e torná-lo parte importante de uma prática de revisão de materiais instrucionais no ensino superior. Confiar apenas na opinião de especialistas para decidir sobre a qualidade de livros textos é muito pouco. É preciso instituir práticas sistemáticas de coleta de dados junto aos estudantes que utilizam tais livros.

Fabiana Queiroga*

Psicóloga pela Universidade Federal da Paraíba. Doutoranda Pós-Graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília, Brasília, DF – Brasil.

Jairo Eduardo Borges-Andrade

Doutor em Sistemas Instrucionais pela Florida State University. Professor titular do Departamento de Psicologia Social e do Trabalho/Pós-Graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília, Brasília, DF – Brasil.

E-mail: jairo@unb.br

Rodrigo Araújo de Miranda

Psicólogo pela Universidade de Brasília. Mestrando da Pós-Graduação em Psicologia Social, do Trabalho e das Organizações da Universidade de Brasília, Brasília, DF – Brasil.

E-mail: rodrigo.miranda2710@gmail.com

***Endereço para envio de correspondência:**

SHIN CA5 Bloco H Apto 512 - Lago Norte, Brasília, DF - Brasil - CEP: 71503-505

E-mail: fabiana.queiroga@gmail.com

Recebido 19/10/2008, Reformulado 02/05/2009, Aprovado 11/05/2009

Referências

- Bell, B., & Cowie, B. (2001a). *Formative assesment and science education*. Norwell, MA: Kluwer Academic Publishers.
- Bell, B., & Cowie, B. (2001b). The characteristics of formative assessment in science education. *Science Education*, 85, 536-553.
- Borges-Andrade, J. E. (2006). Avaliação integrada e somativa em TD&E. In J. E. Borges-Andrade, G. S. Abbad, & L. Mourão (Orgs.), *Treinamento, desenvolvimento e educação em organizações e trabalho: fundamentos para a gestão de pessoas* (pp. 343-358). Porto Alegre: Artmed.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assesment, Research and Evaluation*, 10(7). Recuperado em 29 de novembro de 2006, de <http://pareonline.net/getvn.asp?v=10&n=7>
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Dick, W., & Carey, L. (1978). *The systematic design of instruction*. Chicago: Scott, Foresman and Company.
- Enzmann, D. (2003). *RanEgen 2.0*. Recuperado em 29 de novembro de 2006, de <http://www.kfn.de/softwareenzmann.html>
- Gagné, R. M. (1988). *Essentials of learning for instruction*. Englewood Cliffs, NJ: Prentice Hall.
- Gagné, R. M. (1987). *Instructional tecnology: Foundations*. New Jersey: Laurence Erlbaum Associates.
- Gagné, R. M., & Briggs, L. J. (1974). *Principles of instructional designs*. New York: Holt, Rinehart and Winston.
- Gorsuch, R. L. (1973). Using Barlett's significance test to determine the number of factors to extract. *Educational and Psychological Measurement*, 33, 361-364.
- Jonassen, D. H., Grabinger, R. S., & Harris, N. D. C. (1990). Analyzing instructional strategies and tatics. *Performance and Instructional Quarterly*, 3, 29-45.
- Laros, J. A. (2005). O uso da análise fatorial: algumas diretrizes para pesquisadores. In L. Pasquali (Org.), *Análise fatorial para pesquisadores* (pp. 163-184). Brasília, DF: LabPAM.
- Miles, J., & Shevlin, M. (2001). *Applying regression & correlations – A guide for students and researchers*. London: SAGE.
- Otsuka, J. L., & Rocha, H. V. (2002). Avaliação formativa em ambientes de EaD. *Anais do XIII Simpósio Brasileiro de Informática na Educação*, 1-10. Recuperado em 02 de maio de 2009, de http://www.teleduc.org.br/artigos/17_jh_sbie2002.pdf
- Pantoja, M. J., & Borges-Andrade, J. E. (2009). Estratégias de aprendizagem no trabalho em diferentes ocupações profissionais. *RAC Eletrônica* [Online], 3, 41-62.
- Pasquali, L. (2005). *Análise fatorial para pesquisadores*. Brasília, DF: LabPAM.
- Queiroga, F., Borges-Andrade, J. E., Mauro, T. G. da S. (2004). Uma experiência de avaliação formativa em ensino de psicologia organizacional e do trabalho. In *Resumos de Comunicações Científicas, I Congresso Brasileiro de Psicologia Organizacional e do Trabalho*. Salvador: Sociedade Brasileira de Psicologia Organizacional e do Trabalho.
- Scriven, M., Tyler, R., & Gagné, R. (1967). *Perspectives of curriculum evaluation*. Chicago: Rand McNally.
- Tabachnick, B., & Fidell, L. S. (2001). *Using multivariate statistics*. San Francisco: Allyn & Bacon.
- Zanelli, J. C., Borges-Andrade, J. E., & Bastos, A. V. B. (2004). *Psicologia, organizações e trabalho no Brasil* (576 pp.). Porto Alegre: Artmed.
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 3, 432-442.